

Testimony of Douglas Holtz-Eakin  
Director, Maurice R. Greenberg Center for Geoeconomic Studies  
Paul A. Volcker Chair in International Economics  
Council on Foreign Relations

before

Committee on the Budget  
United States House of Representatives  
September 13, 2006

Chairman Nussle, Ranking Member Spratt and members of the Committee, I am pleased to have the opportunity to discuss the topic of dynamic scoring in the federal budget process. In my remarks, I wish to make observations that fall into three broad areas:

- The principle of dynamic scoring is good science that would potentially bring into the budget process greater information regarding beneficial economic policies,
- Dynamic scoring faces difficulties of implementation in the budget process, and
- Dynamic scoring is not a panacea for either policymaking or the budget process.

Let me cover each in turn before taking your questions.

### **Dynamic Scoring is Good Science**

Budget “scores” are estimates of the change in the federal unified budget that would result from the passage of specific statutory language. All proposals are measured relative to a single, fixed baseline outlook for the budget which is, in turn, built upon a projection for the United States economy. A key feature of current scoring is that in evaluating legislation, the *aggregate* amount of economic activity – total production and income – is assumed to be unchanged from its baseline values.

It is this feature that has led some observers to refer to current scoring procedures as “static.” Unfortunately, this label has caused certain critics to mistakenly conclude that current procedures do not recognize the incentive effects of legislation – that firms, workers, investors, and households continue their economic lives as if nothing had changed. Nothing could be further from the truth. For example, in scoring the impact of the Medicare Modernization Act (MMA), congressional analysts necessarily had to incorporate the decision of firms to offer insurance contracts for the cost of outpatient pharmaceuticals and bid for customers, the willingness of seniors to purchase such insurance, changes in the amount of drugs prescribed and purchased, take-up of low-income subsidies, and a myriad other decisions by households, firms, and governments.

However, in keeping with current practice, the overall level of gross domestic product and national income was assumed to be unchanged.

Dynamic scoring would expand the range of economic impacts to include the pace of economic growth – that is, estimating the change in the aggregate level of economic output and income. This has some desirable features. In estimating the impact of the legislation, analysts would (a) consider the direct impacts on program costs and tax receipts; (b) evaluate the effects on incentives to work, save, invest and conduct economic affairs; (c) estimate the resulting change in the overall level of economic activity; (d) compute the impact of this higher or lower level of economic activity on program costs and tax receipts; and (e) calculate the net impact of the legislation on the unified budget. The key difference is step (d), which is in turn built upon (c).

A virtue of dynamic scoring is that it extends analysis of budget policy to include economic policy dimensions. Specifically, dynamic scoring requires that analysts incorporate into their evaluation of legislation *all* the economic feedbacks at the individual, household, firm, and national level. For this reason, it has the potential to distinguish between those policies which are equal in their budget cost, but very different in their economic incentives. Indeed, one of the most attractive aspects of dynamic scoring is its promise of allowing policymakers to distinguish between economically efficient tax and spending policies that promote growth, and those that work to reduce the living standards of future generations.

### **Difficulties in the Practice of Dynamic Scoring**

The mechanics of doing dynamic scoring are not new. Indeed, a dynamic score can be thought of as the difference between two full-blown baseline budget projections: one in the absence of the legislation, and one in the presence of the proposed legislation. But the scale of the analysis involved in preparing baseline budget projections points to the first problem with wholesale adoption of dynamic scoring: time. In many, if not most, instances statutory language continues to evolve throughout the legislative process: committee deliberation and reporting, floor amendments and votes, and conference committee negotiations. Often there is a need for very quick and timely scoring information. The scale of a dynamic scoring effort may be in conflict with this need.

A second practical difficulty with dynamic scoring is the need for a single, consensus estimate. The attraction of dynamic scoring is its ability to reveal the impact of legislation on economic growth. However, this impact depends crucially on the overall foresightedness of U.S. households and firms. To take an extreme case, imagine legislation that cuts all marginal tax rates by five percentage points, with the cut to take effect five years from now, but sunset ten years in the future. If people are extremely myopic, this policy has no impact on incentives to work, save or invest and there is no dynamic feedback. If they are moderately forward-looking, they may anticipate lower taxes and respond to these incentives. If they are even more forward-looking, they will recognize both the tax reduction and the subsequent rise. As a result, they will work

especially hard during the intervening years – yielding a larger increase in output, incomes, and taxes – with a sharper decline when taxes rise again.

One approach to this problem, exemplified by the Congressional Budget Offices macroeconomic analysis of the president’s budget proposals, is to provide a variety of estimates, each corresponding to a different degree of foresight. However, the budget scoring process would require a single set of estimates, implying that a consensus be reached on a wide variety of issues of this type: foresightedness, the pace of international capital flows, saving responses of households and firms, and so forth.

The example sketched above highlights another issue in the conduct of dynamic scoring: the need for a standard “offsetting policy.” Over the long-term, if individuals have foresight then government debt (relative to the economy) must stabilize. Legislative proposals that upset this requirement by increasing spending or reducing taxes (at least relative to their impact on economic growth) will produce debt that will grow explosively. Similarly, spending cuts or tax increases (relative to their impact on the economy) will cause debt to spiral down. Since the government can neither borrow nor save unboundedly large amounts, it is necessary to put a stop to either spiral by introducing an offsetting budget policy at some point in the future.

The choice of policy – spending increases or decreases and the pace at which they take place, tax reductions or increases and their timing, or some combination of these – will affect the behavior of individuals and firms and influence the score. Since a primary objective of scoring is to treat all legislative proposals equally, it will be necessary to pick a single type of offsetting policy and use it for all proposals.

Another challenge in implementing dynamic scoring is the degree to which the score reflects only supply-side growth, or also includes demand-side cyclical influences. Broadly speaking, economies grow in one of two ways. Supply-side growth occurs when there is an increase in the capacity to produce goods and services through the addition of greater labor supply (labor force participation, hours worked, higher effort per hour, greater skills per worker, better efficiency in the use of labor effort and skills, and so forth), greater physical capital (more or better equipment, software, buildings, and so forth) and improved technical prowess (new technologies or superior organization and management).

Demand-side growth (or contraction) reflects business cycle fluctuations in the extent to which *existing* labor supply, capital, and technical prowess are utilized. The attention paid to monetary and other stabilization policies is clear tribute to the fact that recessions are costly and faster recoveries are desirable. But these changes are transitory and it may not be desirable to include transitory components in the budgetary evaluation of legislative changes.

If these effects are included, they will depend crucially on whether the budget baseline projection begins in a period of recession or boom. If it is the former, then positive

demand effects will augment growth. If it is the latter, growth is limited and the result will be faster onset of return to supply-side potential and greater inflationary pressures.

Finally, the ultimate size, direction, and character of demand-side effects depend as well upon the assumed path of monetary policy. In a manner similar to offsetting budget policies, it would be necessary to make assumptions regarding the response of monetary policy to the legislative changes.

A final issue that arises in full-blown use of dynamic scoring is the interaction between taxes and spending. At present, the Congressional Budget Office scores spending proposals and the Joint Committee on Taxation scores the bulk of tax legislation. By its nature, dynamic scoring seeks to identify the indirect spending consequences of tax legislation and *vice versa*. Accordingly, it will be necessary for these groups to coordinate extensively their respective efforts.

### **Dynamic Scoring is not a Panacea**

One occasionally hears that dynamic scoring is desirable because it will be more accurate. While dynamic scoring will more fully incorporate a wider range of behavioral responses, it is not likely to improve accuracy. First, the mechanical nature of scoring – evaluating different policy proposals using a baseline fixed at the beginning of the legislative calendar – is necessary for even-handed evaluation of alternative proposals, but hardly a recipe for improved accuracy in an ever-changing economy. Further, as noted earlier, the same level, legislative playing field necessarily entails identical and “unrealistic” assumptions regarding offsetting budget policies and monetary policy. Finally, to the extent that the pursuit of good policy leads to a decision to focus on long-run, supply-side growth then the elimination of cycles moves scoring even further away from “accurate” predictions.

Similarly, any move to dynamic scoring would not eliminate the need for analysts making judgment decisions. Quite the contrary, as noted above, in addition to the plethora of issues that already exist (*e.g.*, how fast will legislation become law; how quickly will administrative rule-making be completed; what will implementing regulation look like; how fast will awareness spread and program participation rise?) additional decisions will be needed on the nature of economic growth policies’ ability to influence it.

Neither of these outcomes is bad. The combination of baseline projections and budget scores is intended to support the legislative process, not forecast the economy. There are far more parsimonious and accurate forecasting procedures available. Evaluating innovative legislative proposals necessarily requires analytic judgment because there is literally no policy track record on which to rely. Dynamic scoring may reflect a change in the desired content of the budget process; it does not change the fact that scoring supports that process.

Finally, the greatest reason that dynamic scoring is not a panacea is that it is unlikely to change the bottom line very much. The entire federal budget is only one-fifth the U.S. economy, and few legislative proposals affect even a fraction of the outlay or receipts stream. That is, most legislative proposals don't have enough overall "bang" to generate much dynamics. Of course, some have superior incentive effects – a big "bang for the buck." But even the dynamics of these proposals are not likely to look very large. Over the period from 1820 to 1998, output per capita in the United States grew an average of 0.4 percentage points faster than in the United Kingdom (1.74 versus 1.35 percent per year). Thus, 0.4 percentage points per year – which transformed the global economic order – is a *big* supply-side growth-effect.

For this reason, some have proposed restricting dynamic scoring to particularly comprehensive tax or spending proposals such as tax or social security reform. While sensible in itself, taken at face value it would produce an asymmetry between proposals evaluated with traditional scoring and those that were evaluated using dynamic scoring.

### **Conclusion**

Mr. Chairman and members of the Committee, dynamic scoring is an important and potentially valuable tool for Congress to use in evaluating legislative proposals. I am grateful to have the opportunity to discuss my views on the issue, and look forward to answering your questions.